(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 November 2004 (11.11.2004)

PCT

(10) International Publication Number
WO 2004/097052 A2

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number:
PCT/US2004/013587

(22) International Filing Date: 29 April 2004 (29.04.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/466,067 29 April 2003 (29.04.2003) US
60/538,246 23 January 2004 (23.01.2004) US

(71) Applicant *(for all designated States except US)*: WYETH [US/US]; 5 Giralda Farms, Madison, NJ 07940 (US).

(71) Applicant and
(72) Inventor: STRAHS, Andrew [US/US]; 30 McKinley Street, Maynard, MA 01754 (US).

(72) Inventor: TREPICCHIO, William, L.; 21 Abbott Ridge Drive, Andover, MA 01810 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BURCZYNSKI, Michael, E. [US/US]; 71 Franklin Avenue, Swampscott, MA 01907 (US). TWINE, Natalie, C. [US/US]; 379 Shirley Hill Road, Goffstown, NH 03045 (US). SLONIM, Donna, K. [US/US]; 799 Dale Street, North Andover, MA 01845 (US). IMMERMAN, Fred [US/US]; 194 Haverstraw Road, Suffern, NY 10901 (US). DORNER, Andrew, J. [US/US]; 20 Baskin Road, Lexington, MA 02421 (US).

(74) Agent: VAN DYKE, Raymond; Nixon Peabody LLP, 401 9th Street, N.W., Washington, DC 20004 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS FOR PROGNOSIS AND TREATMENT OF SOLID TUMORS

(57) Abstract: Solid tumor prognosis genes, and methods, systems and equipment of using these genes for the prognosis and treatment of solid tumors. Prognosis genes for a solid tumor can be identified by the present invention. The expression profiles of these genes in peripheral blood mononuclear cells (PBMCs) are correlated with clinical outcome of the solid tumor. The prognosis genes of the present invention can be used as surrogate markers for predicting clinical outcome of a solid tumor in a patient of interest. These genes can also be used to select a treatment which has a favorable prognosis for the solid tumor of the patient of interest.

WO 2004/097052 A2

METHODS FOR PROGNOSIS AND TREATMENT OF SOLID TUMORS

[0001] The present invention incorporates by reference all materials recorded in the compact discs labeled "Copy 1 – Sequence Listing Part" "Copy 2 – Sequence Listing Part" and "Copy 3 – Sequence Listing Part," each of which includes "Sequence Listing.ST25.txt" (5,454 KB, created April 28, 2004). The present invention also incorporates by reference all materials recorded in the compact discs labeled "Copy 1 – Tables Part," "Copy 2 – Tables Part," and "Copy 3 – Tables Part," each of which includes the following files: "Table 3 - Spearman Correlation of Baseline Expression with Clinical Outcome.txt" (298 KB, created April 28, 2004), "Table 4 - Qualifiers and the Corresponding Entrez and Unigene Accession Nos.txt" (179 KB, created April 28, 2004), "Table 5 - Genes and Gene Titles.txt" (331 KB, created April 28, 2004), and "Table 8 - Cox Regression of Clinical Outcome on Baseline Gene Expression.txt" (294 KB, created April 28, 2004).

CROSS-REFERENCE TO RELATED APPLICATIONS

[0002] The present application claims priority from and incorporates by reference the entire disclosures of U.S. Provisional Patent Application Serial No. 60/466,067, filed April 29, 2003, and U.S. Provisional Patent Application Serial No. 60/538,246, filed January 23, 2004.

TECHNICAL FIELD

[0003] The present invention relates to solid tumor prognosis genes and methods of using these genes for the prognosis or treatment of solid tumors.

BACKGROUND

[0004] Expression profiling studies in primary tissues have demonstrated that there exist transcriptional differences between normal and malignant tissues. See, for example, Su, *et al.*, CANCER RES, 61:7388-7393 (2001); and Ramaswamy, *et al.*, PROC NATL ACAD SCI U.S.A., 98:15149-15151 (2001). Recent clinical analyses have also identified expression profiles within tumors that appear to be highly correlated with certain measures of clinical outcomes. One study has demonstrated that expression profiling of primary tumor biopsies yields prognostic "signatures" that rival or may even out-perform currently

accepted standard measures of risk in cancer patients. See van de Vijver, *et al.*, N ENGL J MED, 347:1999-2009 (2002).

SUMMARY OF THE INVENTION

[0005] The present invention provides methods, systems and equipment for prognosis or selection of treatment of solid tumors. Prognosis genes for a solid tumor can be identified by the present invention. The expression profiles of these genes in peripheral blood mononuclear cells (PBMCs) are correlated with clinical outcome of the solid tumor. These genes can be used as surrogate markers for predicting clinical outcome of the solid tumor in a patient of interest. These genes can also be used to identify or select treatments which have favorable prognoses for the patient of interest.

[0006] In one aspect, the present invention provides methods that are useful for the prognosis or selection of treatment of a solid tumor in a patient of interest. The methods include comparing an expression profile of one or more prognosis genes in a peripheral blood sample of the patient of interest to at least one reference expression profile of the prognosis genes. Each of the prognosis genes is differentially expressed in PBMCs of a first class of patients as compared to PBMCs of a second class of patients. Both classes of patients have a solid tumor, and each class of patients has a different clinical outcome. In many embodiments, the prognosis genes are substantially correlated with a class distinction between the two classes of patients.

[0007] Solid tumors amenable to the present invention include, but are not limited to, renal cell carcinoma (RCC), prostate cancer, head/neck cancer, and other tumors that do not have their origin in blood or lymph cells.

[0008] Clinical outcome can be measured by any clinical indicator. In one embodiment, clinical outcome is determined based on clinical classifications such as complete response, partial response, minor response, stable disease, progressive disease, non-progressive disease, or any combination thereof. In another embodiment, clinical outcome is measured by time to disease progression (TTP) or time to death (TTD). In still another embodiment, clinical outcome is prognosticated by using traditional risk assessment methods, such as Motzer risk classification for RCC. Other patient responses to a therapeutic treatment can also be used to measure clinical outcome. Examples of solid tumor treatments include, but are not limited to, drug therapy (e.g., CCI-779 therapy),

chemotherapy, hormone therapy, radiotherapy, immunotherapy, surgery, gene therapy, anti-angiogenesis therapy, palliative therapy, or any combination thereof.

[0009] In many embodiments, the reference expression profile(s) includes an average expression profile of the prognosis genes in peripheral blood samples of reference patients. In many instances, the reference patients have the same solid tumor as the patient of interest, and the clinical outcome of the reference patients are either known or determinable.

[0010] The peripheral blood samples of the patient of interest and reference patients can be whole blood samples, or blood samples comprising enriched or purified PBMCs. Other types of blood samples can also be employed in the present invention. In one embodiment, all of the peripheral blood samples are baseline samples which are isolated from respective patients prior to a therapeutic treatment of the patients.

[0011] Any comparison method can be used to compare the expression profile of the patient of interest to the reference expression profile(s). In one embodiment, the comparison is based on the absolute or relative peripheral blood expression level of each prognosis gene. In another embodiment, the comparison is based on the ratios between expression levels of two or more prognosis genes. In yet another embodiment, the reference expression profiles include at least two distinct expression profiles, each being derived from a different class of reference patients. The comparison of the expression profile of the patient of interest to the reference expression profiles can be carried out by using methods including, but not limited to, hierarchical clustering, *k*-nearest-neighbors, or weighted-voting algorithm.

[0012] In still another embodiment, the methods of the present invention include selecting a treatment which has a favorable prognosis for the solid tumor in the patient of interest.

[0013] In another aspect, the present invention provides other methods useful for the prognosis or selection of treatment of a solid tumor in a patient of interest. These methods include comparing an expression profile of one or more prognosis genes in a peripheral blood sample of the patient of interest to at least one reference expression profile of the prognosis genes, where each of the prognosis genes is differentially expressed in PBMCs of a first class of patients as compared to PBMCs of a second class of patients. Each of the first and second classes is a subcluster formed by an unsupervised clustering analysis of gene expression profiles in PBMCs of patients who have the solid tumor. In one

embodiment, the majority of the first class of patients has a first clinical outcome, and the majority of the second class of patients has a second clinical outcome.

[0014] In yet another aspect, the present invention further provides methods useful for the prognosis or selection of treatment of a solid tumor in a patient of interest. The methods include comparing an expression profile of one or more prognosis genes in a peripheral blood sample of the patient of interest to at least one reference expression profile of the prognosis genes, where the expression levels of each of the prognosis genes in PBMCs of patients having the solid tumor are correlated with clinical outcomes of these patients. The association between PBMC expression levels and clinical outcome can be determined by a statistical method (e.g., Spearman's rank correlation or Cox proportional hazard regression model) or a class-based correlation metric (e.g., neighborhood analysis). In one embodiment, the solid tumor is RCC, and clinical outcome is measured by patient response to a CCI-779 therapy. In another embodiment, the prognosis genes include at least one gene selected from Tables 6a, 6b, 6c, 6d, 9a, 9b, 9c, 9d, 10, 11, 12, 13, 16, 20, and 21.

[0015] The present invention also features systems useful for the prognosis or selection of treatment of a solid tumor in a patient of interest. The systems include (1) a memory or a storage medium comprising data that represent an expression profile of one or more prognosis genes in a peripheral blood sample of the patient of interest, (2) a storage medium comprising data that represent at least one reference expression profile of the prognosis genes, (3) a program capable of comparing the expression profile of the patient of interest to the reference expression profile, and (4) a processor capable of executing the program. The expression levels of the prognosis genes in PBMCs of patients having the solid tumor are correlated with clinical outcomes of the patients.

[0016] Moreover, the present invention features nucleic acid or protein arrays useful for the prognosis or selection of treatment of a solid tumor in a patient of interest. The nucleic acid or protein arrays include concentrated probes for solid tumor prognosis genes.

[0017] Other features, objects, and advantages of the present invention are apparent in the detailed description that follows. It should be understood, however, that the detailed description, while indicating embodiments of the present invention, is given by way of illustration only, not limitation. Various changes and modifications within the scope of the invention will become apparent to those skilled in the art from the detailed description.

BRIEF DESCRIPTION OF THE DRAWINGS

[0018] The drawings are provided for illustration, not limitation. All drawings in the parallel U.S. patent application, entitled "Methods for Prognosis and Treatment of Solid Tumors" and filed April 29, 2004, are incorporated herein by reference.

[0019] Figure 1A depicts expression profiles of class-correlated genes identified by nearest-neighbor analysis of patients with survival of less than 150 days versus patients with survival of greater than 550 days. The relative expression levels of the class-correlated genes (rows) are indicated for each patient (columns) according to the normalized expression level scale.

[0020] Figure 1B shows the comparison of the signal to noise (S2N) similarity metric scores for class-correlated genes identified in Figure 1A relative to S2N scores for the top 1%, 5%, and 50% of scores for class-correlated genes resulting from randomly permuted data sets.

[0021] Figure 1C illustrates training set cross validation results for predictor gene sets of increasing size. Each predictor set was evaluated by cross validation to identify the predictor set with the highest accuracy for classification of the samples. In these analyses, a 58 gene predictor set (77% accuracy) was the optimal classifier.

[0022] Figure 1D demonstrates cross validation results for each sample using the 58-gene predictor identified in Figure 1C. A leave-one-out cross validation was performed and the prediction strengths were calculated for each sample in the analysis. For the purposes of illustration, confidence scores accompanying calls of "TTD > 550 days" were assigned positive values, while prediction strengths accompanying calls of "TTD < 150 days" were assigned negative values.

[0023] Figure 2A shows the relative gene expression levels of a 42-gene classifier for the comparison of patients with intermediate versus poor Motzer risk classification.

[0024] Figure 2B shows the relative gene expression levels for an 18-gene classifier identified in the comparison of patients with progressive disease versus any other clinical response.

[0025] Figure 2C demonstrates the relative gene expression levels for a 6-gene classifier identified in the comparison of patients in the lower versus upper quartiles of time to disease progression.

[0026] Figure 2D shows the relative gene expression levels for a 52-gene classifier identified in the comparison of patients in the lower versus upper quartiles of survival/time to death.

[0027] Figure 2E depicts the relative expression levels for a 12-gene classifier identified in the comparison of patients with early (time to disease progression < 106 days) versus all other times to disease progression (TTP ≥ 106 days).

[0028] Figure 3A illustrates the dendrogram of an unsupervised hierarchical clustering of baseline PBMC profiles in 45 RCC patients using all expressed genes present in at least one sample and possessing a frequency of greater than 10 ppm in at least one sample (5,424 genes total). PBMC expression profiles in the poor prognosis cluster are indicated by subcluster "A," where 9 out of 12 patients with PBMC profiles in this subcluster exhibited survival of less than a year. PBMC expression profiles in the good prognosis cluster are indicated by subcluster "C," where 10 out of 12 patients with PBMC profiles in this subcluster exhibited survival of greater than a year. The median survival for patients in subclusters A, B, C, and D is 281 days, 566 days, 573 days, and 502 days, respectively.

[0029] Figure 3B shows baseline expression profiles of selected genes in RCC patients. The dendrogram of sample relatedness is indicated.

[0030] Figure 4A illustrates the Kaplan-Meier survival curve for patients in the poor and good prognosis subclusters segregated on the basis of gene expression pattern.

[0031] Figure 4B illustrates the Kaplan-Meier survival curve for patients in the poor and good prognosis subclusters segregated on the basis of Motzer risk assessment.

[0032] Figure 5A demonstrates the result of supervised identification of a gene classifier for assigning class membership to patients in the good and poor prognosis subclusters. The relative expression levels of the most class-correlated gene (rows) are indicated for each patient (columns) according to the scale described in Figure 1A.

[0033] Figure 5B shows cross validation results for each sample using the gene classifier of Figure 5A. A leave-one-out cross validation was performed and the confidence scores were calculated for each sample in the analysis. Similar to Figure 1D, for the purposes of illustration, prediction strengths accompanying calls of "survival > 1 year" were assigned positive values, while prediction strengths accompanying calls of "survival < 1 year" were assigned negative values. Asterisks identify the false positives in this clinical assay designed to identify short survival times, and arrowheads indicate false negatives.

[0034] Figure 6A shows the optimal gene classifier for year-long survival identified by nearest-neighbor analysis using a more stringent filter (at least 25% present calls, and an average frequency no less than 5 ppm). A GeneCluster gene selection approach identifies

genes distinguishing patients with survival less than 365 days versus patients with survival greater than 365 days in the training set. The relative expression levels of the most class-correlated genes (rows) are indicated for each of the patients in the training set (columns) according to the scale described in Figure 1A.

[0035] Figure 6B evaluates prediction accuracy of gene classifiers of increasing size. Accuracy of class assignment for gene classifiers containing between 2 and 60 genes in steps of 2, and 60-200 genes in steps of 10, were evaluated by leave-one-out cross validation on the training set of samples. The smallest predictive model with the highest accuracy was selected (20 gene predictor, indicated by the arrow).

[0036] Figure 6C demonstrates the result of evaluation of the optimal predictive model of Figure 6B on an untested set of RCC PBMC profiles. A *k*-nearest-neighbors algorithm using the 20 gene classifier was used to assign class membership to the remaining 14 PBMC profiles, and the prediction strengths associated with the class assignments are presented for each sample in the analysis. For the purposes of illustration, confidence scores accompanying calls of "TTD < 365 days" were assigned positive values, while confidence scores accompanying calls of "TTD > 365 days" were assigned negative values. The overall accuracy of the gene classifier was 72%. By defining the clinical assay as the identification of favorable outcome, eight of eight patients with favorable outcome were correctly identified as having survival greater than one year (positive predictive value of 100%).

[0037] Figure 7A illustrates the optimal gene classifier for greater than 106 day time to progression identified by nearest-neighbor analysis using a more stringent filter (at least 25% present calls, and an average frequency no less than 5 ppm). A GeneCluster gene selection approach identifies genes distinguishing patients with TTP less than 106 days versus patients with TTP greater than 106 days in the training set. The relative expression levels of the most class-correlated genes (rows) are indicated for each of the patients in the training set (columns) according to the scale of Figure 1A.

[0038] Figure 7B indicates prediction accuracy of gene classifiers of increasing size. Accuracy of class assignment for gene classifiers containing between 2 and 60 genes in steps of 2, and 60-200 genes in steps of 10, were evaluated by leave-one-out cross validation on the training set of samples. The smallest predictive model with the highest accuracy was selected (30 gene predictor, indicated by the arrow).

[0039] Figure 7C shows the result of evaluation of the optimal predictive model of Figure 7B on an untested set of RCC PBMC profiles. A *k*-nearest-neighbors algorithm using the 30 gene classifier was used to assign class membership to the remaining 14 PBMC profiles, and the prediction strengths associated with the class assignments are presented for each sample in the analysis. For the purposes of illustration, confidence scores accompanying calls of "TTP < 106 days" were assigned positive values, while confidence scores accompanying calls of "TTD > 106 days" were assigned negative values. The overall accuracy of the gene classifier was 85%. By defining the clinical assay as the identification of favorable outcome, eight of ten patients with favorable outcome were correctly identified as having TTP greater than one 106 days (positive predictive value of 80%) and three of three patients with poor outcome were correctly predicted to have TTP less than 106 days (negative predictive value 100%).

DETAILED DESCRIPTION

[0040] The present invention provides methods that are useful for prognosis or selection of treatment of solid tumors. These methods employ prognosis genes that are differentially expressed in peripheral blood samples of solid tumor patients who have different clinical outcomes. In many embodiments, the peripheral blood expression profiles of these prognosis genes are correlated with patients' clinical outcome or prognosis under a statistical method or a correlation model. In many other embodiments, solid tumor patients can be divided into at least two classes based on patients' clinical outcome or prognosis, and the prognosis genes are substantially correlated with a class distinction between these two classes of patients under a neighborhood analysis.

[0041] The prognosis genes of the present invention can be used as surrogate markers for the prediction of clinical outcome of solid tumors. The prognosis genes of the present invention can also be used for the identification of optimal treatments of solid tumors. Different patients may have distinct clinical responses to a therapeutic treatment due to individual heterogeneity of the molecular mechanism of the disease. The identification of gene expression patterns that correlate with patient response allows clinicians to select treatments based on predicted patient responses and thereby avoid adverse reactions. This provides improved power and safety of clinical trials and increased benefit/risk ratio for drugs and other therapeutic treatments. Peripheral blood is a tissue that can be routinely obtained from patients in a minimally invasive manner. By determining the correlation

between patient outcome and gene expression profiles in peripheral blood samples, the present invention represents a significant advance in clinical pharmacogenomics and solid tumor treatment.

[0042] Various aspects of the invention are described in further detail in the following subsections. The use of subsections is not meant to limit the invention. Each subsection may apply to any aspect of the invention. In this application, the use of "or" means "and/or" unless stated otherwise.

I. General Methods for Identifying Solid Tumor Prognosis Genes

[0043] Previous studies demonstrated that baseline expression profiles in PBMCs from solid tumor patients were significantly distinct from those of disease-free subjects. See U.S. Provisional Application Serial No. 60/459,782, filed April 3, 2003, U.S. Provisional Application Serial No. 60/427,982, filed November 21, 2002, and U.S. Patent Application Serial No. 10/717,597, filed November 21, 2003, all of which are incorporated herein by reference. Studies also showed that gene expression profiles in PBMCs were predictive of anti-cancer drug activity *in vivo*. See U.S. Provisional Application Serial No. 60/446,133, filed February 11, 2003, and U.S. Patent Application Serial No. 10/775,169, filed February 11, 2004, both of which are incorporated herein by reference. In addition, studies indicated that PBMC baseline expression profiles were correlated with clinical outcomes of RCC or other non-blood diseases. See U.S. Provisional Application Serial No. 60/466,067, filed April 29, 2003, which is incorporated herein by reference.

[0044] The present invention further evaluates the correlation between peripheral blood gene expression and clinical outcome of solid tumors. Prognosis genes for a variety of solid tumors can be identified by the present invention. These genes are differentially expressed in peripheral blood samples of solid tumor patients who have different clinical outcomes. In many embodiments, the peripheral blood expression profiles of the prognosis genes of the present invention are correlated with patient outcome under statistical methods or correlation models. Exemplary statistical methods and correlation models include, but are not limited to, Spearman's rank correlation, Cox proportional hazard regression model, ANOVA/t test, nearest-neighbor analysis, and other rank tests, survival models or class-based correlation metrics.

[0045] Solid tumors amenable to the present invention include, without limitation, RCC, prostate cancer, head/neck cancer, ovarian cancer, testicular cancer, brain tumor, breast cancer, lung cancer, colon cancer, pancreas cancer, stomach cancer, bladder cancer, skin cancer, cervical cancer, uterine cancer, and liver cancer. In one embodiment, the solid tumors do not have their origin in blood or lymph (hematopoetic) cells. Solid tumors can be measured or evaluated using direct or indirect visualization procedures. Suitable visualization methods include, but are not limited to, scans (such as X-rays, computerized axial tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), or ultrasonography (U/S)), biopsy, palpation, endoscopy, laparoscopy, and other suitable means as appreciated by those skilled in the art.

[0046] Clinical outcome of solid tumors can be assessed by numerous criteria. In many embodiments, clinical outcome is assessed based on patients' response to a therapeutic treatment. Examples of clinical outcome measures include, without limitation, complete response, partial response, minor response, stable disease, progressive disease, time to disease progression (TTP), time to death (TTD or Survival), or any combination thereof. Examples of solid tumor treatments include, without limitation, drug therapy (e.g., CCI-779 therapy), chemotherapy, hormone therapy, radiotherapy, immunotherapy, surgery, gene therapy, anti-angiogenesis therapy, palliative therapy, or any combination thereof, or other conventional or non-conventional therapies.

[0047] In one embodiment, clinical outcome is evaluated based on the WHO Reporting Criteria, such as those described in WHO Publication, No. 48 (World Health Organization, Geneva, Switzerland, 1979). Under the Criteria, uni- or bidimensionally measurable lesions are measured at each assessment. When multiple lesions are present in any organ, up to 6 representative lesions can be selected, if available.

[0048] In another embodiment, clinical outcome is determined based on a classification system composed of clinical categories such as complete response, partial response, minor response, stable disease, progressive disease, or any combination thereof. "Complete response" (CR) means complete disappearance of all measurable and evaluable disease, determined by two observations not less than 4 weeks apart. There is no new lesion and no disease related symptom. "Partial response" (PR) in reference to bidimensionally measurable disease means decrease by at least about 50% of the sum of the products of the largest perpendicular diameters of all measurable lesions as determined by 2 observations not less than 4 weeks apart. "Partial response" in reference to unidimensionally measurable

disease means decrease by at least about 50% in the sum of the largest diameters of all lesions as determined by 2 observations not less than 4 weeks apart. It is not necessary for all lesions to have regressed to qualify for partial response, but no lesion should have progressed and no new lesion should appear. The assessment should be objective. "Minor response" in reference to bidimensionally measurable disease means about 25% or greater decrease but less than about 50% decrease in the sum of the products of the largest perpendicular diameters of all measurable lesions. "Minor response" in reference to unidimensionally measurable disease means decrease by at least about 25% but less than about 50% in the sum of the largest diameters of all lesions.

[0049] "Stable disease" (SD) in reference to bidimensionally measurable disease means less than about 25% decrease or less than about 25% increase in the sum of the products of the largest perpendicular diameters of all measurable lesions. "Stable disease" in reference to unidimensionally measurable disease means less than about 25% decrease or less than about 25% increase in the sum of the diameters of all lesions. No new lesions should appear. "Progressive disease" (PD) refers to a greater than or equal to about a 25% increase in the size of at least one bidimensionally (product of the largest perpendicular diameters) or unidimensionally measurable lesion or appearance of a new lesion. The occurrence of pleural effusion or ascites is also considered as progressive disease if this is substantiated by positive cytology. Pathological fracture or collapse of bone is not necessarily evidence of disease progression.

[0050] In yet another embodiment, overall subject tumor response for uni- and bidimensionally measurable disease is determined according to Table 1.

Table 1. Overall Subject Tumor Response

Response in Bidimensionally Measurable Disease	Response in Unidimensionally Measurable Disease	Overall Subject Tumor Response
PD	Any	PD
Any	PD	PD
SD	SD or PR	SD
SD	CR	PR
PR	SD or PR or CR	PR
CR	SD or PR	PR
CR	CR	CR

[0051] Overall subject tumor response for non-measurable disease can be assessed, for instance, in the following situations:

a) Overall complete response: if non-measurable disease is present, it should disappear completely. Otherwise, the subject cannot be considered as an "overall complete responder."

b) Overall progression: in case of a significant increase in the size of non-measurable disease or the appearance of a new lesion, the overall response will be progression.

[0052] Clinical outcome can also be assessed by other criteria. For instance, clinical outcome can be measured by TTP or TTD. TTP refers to the interval from the date of initiation of a therapeutic treatment until the first day of measurement of progressive disease. TTD refers to the interval from the date of initiation of a therapeutic treatment to the time of death, or censored at the last date known alive.

[0053] Moreover, clinical outcome can include prognoses based on traditional clinical risk assessment methods. In many cases, these risk assessment methods employ numerous prognostic factors to classify patients into different prognosis or risk groups. One example is Motzer risk assessment for RCC, as described in Motzer, *et al.*, J CLIN ONCOL, 17:2530-2540 (1999). Patients in different risk groups may have different responses to a therapy.

[0054] Peripheral blood samples employed in the present invention can be isolated from solid tumor patients at any disease or treatment stage. In one embodiment, the peripheral blood samples are isolated from solid tumor patients prior to a therapeutic treatment. These blood samples are "baseline samples" with respect to the therapeutic treatment.

[0055] A variety of peripheral blood samples can be used in the present invention. In one embodiment, the peripheral blood samples are whole blood samples. In another embodiment, the peripheral blood samples comprise enriched PBMCs. By "enriched," it means that the percentage of PBMCs in the sample is higher than that in whole blood. In some cases, the PBMC percentage in an enriched sample is at least 1, 2, 3, 4, 5 or more times higher than that in whole blood. In some other cases, the PBMC percentage in an enriched sample is at least 90%, 95%, 98%, 99%, 99.5%, or more. Blood samples containing enriched PBMCs can be prepared using any method known in the art, such as Ficoll gradients centrifugation or CPTs (cell purification tubes).

[0056] The relationship between peripheral blood gene expression profiles and patient outcome can be evaluated using global gene expression analyses. Methods suitable for this purpose include, but are not limited to, nucleic acid arrays (such as cDNA or oligonucleotide arrays), 2-dimensional SDS-polyacrylamide gel electrophoresis/mass spectrometry, and other high throughput nucleotide or polypeptide detection techniques.

[0057] Nucleic acid arrays allow for quantitative detection of the expression levels of a large number of genes at one time. Examples of nucleic acid arrays include, but are not limited to, Genechip® microarrays from Affymetrix (Santa Clara, CA), cDNA microarrays from Agilent Technologies (Palo Alto, CA), and bead arrays described in U.S. Patent Nos. 6,288,220 and 6,391,562.

[0058] The polynucleotides to be hybridized to nucleic acid arrays can be labeled with one or more labeling moieties to allow for detection of hybridized polynucleotide complexes. The labeling moieties can include compositions that are detectable by spectroscopic, photochemical, biochemical, bioelectronic, immunochemical, electrical, optical or chemical means. Exemplary labeling moieties include radioisotopes, chemiluminescent compounds, labeled binding proteins, heavy metal atoms, spectroscopic markers such as fluorescent markers and dyes, magnetic labels, linked enzymes, mass spectrometry tags, spin labels, electron transfer donors and acceptors, and the like. Unlabeled polynucleotides can also be employed. The polynucleotides can be DNA, RNA, or a modified form thereof.

[0059] Hybridization reactions can be performed in absolute or differential hybridization formats. In the absolute hybridization format, polynucleotides derived from one sample, such as PBMCs from a patient in a selected outcome class, are hybridized to the probes on a nucleic acid array. Signals detected after the formation of hybridization complexes correlate to the polynucleotide levels in the sample. In the differential hybridization format, polynucleotides derived from two biological samples, such as one from a patient in a first outcome class and the other from a patient in a second outcome class, are labeled with different labeling moieties. A mixture of these differently labeled polynucleotides is added to a nucleic acid array. The nucleic acid array is then examined under conditions in which the emissions from the two different labels are individually detectable. In one embodiment, the fluorophores Cy3 and Cy5 (Amersham Pharmacia Biotech, Piscataway N.J.) are used as the labeling moieties for the differential hybridization format.

[0060] Signals gathered from nucleic acid arrays can be analyzed using commercially available software, such as those provide by Affymetrix or Agilent Technologies. Controls, such as for scan sensitivity, probe labeling and cDNA/cRNA quantitation, can be included in the hybridization experiments. In many embodiments, the nucleic acid array expression signals are scaled or normalized before being subject to further analysis. For instance, the expression signals for each gene can be normalized to take into account variations in hybridization intensities when more than one array is used under similar test conditions. Signals for individual polynucleotide complex hybridization can also be normalized using the intensities derived from internal normalization controls contained on each array. In addition, genes with relatively consistent expression levels across the samples can be used to normalize the expression levels of other genes. In one embodiment, the expression levels of the genes are normalized across the samples such that the mean is zero and the standard deviation is one. In another embodiment, the expression data detected by nucleic acid arrays are subject to a variation filter which excludes genes showing minimal or insignificant variation across all samples.

[0061] The gene expression data collected from nucleic acid arrays can be correlated with clinical outcome using a variety of methods. Suitable correlation methods include, but are not limited to, statistical methods (such as Spearman's rank correlation, Cox proportional hazard regression model, ANOVA/t test, or other suitable rank tests or survival models) and class-based correlation metrics (such as nearest-neighbor analysis).

[0062] In one aspect, class-based correlation metrics are used to identify the correlation between peripheral blood gene expression and clinical outcome. In one embodiment, patients with a specified solid tumor are divided into at least two classes based on their clinical stratifications. The correlation between peripheral blood gene expression (e.g., in PBMCs) and clinical outcome is analyzed by a supervised cluster algorithm. Exemplary supervised clustering algorithms include, but are not limited to, nearest-neighbor analysis, support vector machines, and SPLASH. Under the supervised cluster algorithms, clinical outcome of each class of patients is either known or determinable. Genes that are differentially expressed in peripheral blood cells (e.g., PBMCs) of one class of patients relative to the other class of patients can be identified. In many cases, the genes thus identified are substantially correlated with a class distinction between the two classes of patients. The genes thus identified can be used as surrogate markers for predicting clinical outcome of the solid tumor in a patient of interest.

[0063] In another embodiment, patients with a specified solid tumor can be divided into at least two classes based on gene expression profiles in their peripheral blood cells. Methods suitable for this purpose include unsupervised clustering algorithms, such as self-organized maps (SOMs), k-means, principal component analysis, and hierarchical clustering. A substantial number (e.g., at least 50%, 60%, 70%, 80%, 90%, or more) of patients in one class may have a first clinical outcome, and a substantial number of patients in the other class may have a second clinical outcome. Genes that are differentially expressed in the peripheral blood cells of one class of patients relative to the other class of patients can be identified. These genes are prognosis genes for the solid tumor.

[0064] In yet another embodiment, patients with a specified solid tumor can be divided into three or more classes based on their clinical stratifications or peripheral blood gene expression profiles. Multi-class correlation metrics can be employed to identify genes that are differentially expressed in these classes. Exemplary multi-class correlation metrics include, but are not limited to, GeneCluster 2 software provided by MIT Center for Genome Research at Whitehead Institute (Cambridge, MA).

[0065] In a further embodiment, nearest-neighbor analysis (also known as neighborhood analysis) is used to analyze gene expression data gathered from nucleic acid arrays. The algorithm for neighborhood analysis is described in Golub, *et al.*, SCIENCE, 286: 531-537 (1999), Slonim, *et al.*, PROCS. OF THE FOURTH ANNUAL INTERNATIONAL CONFERENCE ON COMPUTATIONAL MOLECULAR BIOLOGY, Tokyo, Japan, April 8 - 11, p263-272 (2000), and U.S. Patent No. 6,647,341, all of which are incorporated herein by reference. Under one form of the neighborhood analysis, the expression profile of each gene can be represented by an expression vector $g = (e_1, e_2, e_3, \ldots, e_n)$, where e_i corresponds to the expression level of gene "g" in the *ith* sample. A class distinction can be represented by an idealized expression pattern $c = (c_1, c_2, c_3, \ldots, c_n)$, where $c_i = 1$ or -1, depending on whether the *ith* sample is isolated from class 0 or class 1. Class 0 may include patients having a first clinical outcome, and class 1 includes patients having a second clinical outcome. Other forms of class distinction can also be employed. Typically, a class distinction represents an idealized expression pattern, where the expression level of a gene is uniformly high for samples in one class and uniformly low for samples in the other class.

[0066] The correlation between gene "g" and the class distinction can be measured by a signal-to-noise score:

$$P(g,c) = [\mu_1(g) - \mu_2(g)]/[\sigma_1(g) + \sigma_2(g)]$$

where $\mu_1(g)$ and $\mu_2(g)$ represent the means of the log-transformed expression levels of gene "g" in class 0 and class 1, respectively, and $\sigma_1(g)$ and $\sigma_2(g)$ represent the standard deviation of the log-transformed expression levels of gene "g" in class 0 and class 1, respectively. A higher absolute value of a signal-to-noise score indicates that the gene is more highly expressed in one class than in the other. In one embodiment, the samples used to derive the signal-to-noise score comprise enriched or purified PBMCs. Thus, the signal-to-noise score P(g,c) can represent a correlation between the class distinction and the expression level of gene "g" in PBMCs.

[0067] The correlation between gene "g" and the class distinction can also be measured by other methods, such as by the Pearson correlation coefficient or the Euclidean distance, as appreciated by those skilled in the art.

[0068] The significance of the correlation between peripheral blood gene expression patterns and the class distinction can be evaluated using a random permutation test. An unusually high density of genes within the neighborhoods of the class distinction, as compared to random patterns, suggests that many genes have expression patterns that are significantly correlated with the class distinction. The correlation between genes and the class distinction can be diagrammatically viewed through a neighborhood analysis plot, in which the y-axis represents the number of genes within various neighborhoods around the class distinction and the x-axis indicates the size of the neighborhood (i.e., P(g,c)). Curves showing different significance levels for the number of genes within corresponding neighborhoods of randomly permuted class distinctions can also be included in the plot.

[0069] In one embodiment, the prognosis genes of the present invention are substantially correlated with a class distinction between two outcome classes. In one example, the prognosis genes are above the median significance level in the neighborhood analysis plot. This means that the correlation measure P(g,c) for each prognosis gene is such that the number of genes within the neighborhood of the class distinction having the size of P(g,c) is greater than the number of genes within the corresponding neighborhoods of randomly permuted class distinctions at the median significance level. In another example, the employed prognosis genes are above the 10%, 5%, 2%, or 1% significance level. As used herein, x% significance level means that x% of random neighborhoods contain as many genes as the real neighborhood around the class distinction.

[0070] Class predictors can be constructed using the prognosis genes of the present invention. These class predictors are useful for assigning class membership to solid tumor

patients. In one embodiment, the prognosis genes in a class predictor are limited to those shown to be significantly correlated with the class distinction by the permutation test, such as those at above the 1%, 2%, 5%, 10%, 20%, 30%, 40%, or 50% significance level. In another embodiment, the expression level of each prognosis gene in a class predictor is substantially higher or substantially lower in PBMCs of one class of patients than in the other class of patients. In still another embodiment, the prognosis genes in a class predictor have top absolute values of P(g,c). In yet another embodiment, the p-value under a Student's *t*-test (e.g., two-tailed distribution, two sample unequal variance) for each differentially expressed prognosis gene is no more than 0.05, 0.01, 0.005, 0.001, 0.0005, 0.0001, or less.

[0071] In a further embodiment, the class predictors of the present invention have at least 50% accuracy for leave-one-out cross validation. In another embodiment, the class predictors of the present invention have at least 60%, 70%, 80%, 90%, 95%, or 99% accuracy for leave-one-out cross validation.

[0072] In another aspect, the correlation between peripheral blood gene expression profiles and clinical outcome can be evaluated by statistical methods. Clinical outcome suitable for these analyses includes, but are not limited to, TTP, TTD, and other time-associated clinical indicators. One exemplary statistical method employs Spearman's rank correlation coefficient, which has the formula of:

$$r_s = SS_{UV}/(SS_{UU}SS_{VV})^{1/2}$$

where $SS_{UV} = \Sigma U_i V_i - [(\Sigma U_i)(\Sigma V_i)]/n$, $SS_{UU} = \Sigma V_i^2 - [(\Sigma V_i)^2]/n$, and $SS_{VV} = \Sigma U_i^2 - [(\Sigma U_i)^2]/n$. U_i is the expression level ranking of a gene of interest, V_i is the ranking of the clinical outcome, and n represents the number of patients. The shortcut formula for Spearman's rank correlation coefficient is $r_s=1 - (6 \times \Sigma d_i^2)/[n(n^2-1)]$, where $d_i = U_i - V_i$. The Spearman's rank correlation is similar to the Pearson's correlation except that it is based on ranks and is thus more suitable for data that is not normally distributed. See, for example, Snedecor and Cochran, STATISTICAL METHODS, Eight edition, Iowa State University Press, Ames, Iowa, 503 pp, 1989. The correlation coefficient is tested to assess whether it differs significantly from a value of 0 (i.e., no correlation).

[0073] The correlation coefficients for each prognosis gene identified by the Spearman's rank correlation can be either positive or negative, provided that the correlation is statistically significant. In many embodiments, the p-value for each prognosis gene thus identified is no more than 0.05, 0.01, 0.005, 0.001, 0.0005, 0.0001, or less. In many other

embodiments, the Spearman correlation coefficients of the prognosis genes thus identified have absolute values of at least 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, or more.

[0074] Another exemplary statistical method is Cox proportional hazard regression model, which has the formula of:

$$\log h_i(t) = \alpha(t) + \beta_j x_{ij}$$

where $h_i(t)$ is the hazard function that assesses the instantaneous risk of demise at time t, conditional on survival to that time, $\alpha(t)$ is the baseline hazard function, and x_{ij} is a covariate which may represent, for example, the expression level of prognosis gene j in a peripheral blood sample. See Cox, JOURNAL OF THE ROYAL STATISTICAL SOCIETY, SERIES B 34:187 (1972). Additional covariates, such as interactions between covariates, can also be included in Cox proportional hazard model. As used herein, the terms "demise" or "survival" are not limited to real death or survival. Instead, these terms should be interpreted broadly to cover any type of time-associated events, such as TTP. In many cases, the p-values for the correlation under Cox proportional hazard regression model are no more than 0.05, 0.01, 0.005, 0.001, 0.0005, 0.0001, or less. The p-values for the prognosis genes identified under Cox proportional hazard regression model can be determined by the likelihood ratio test, Wald test, the Score test, or the log-rank test. In one embodiment, the hazard ratios for the prognosis genes thus identified are at least 1.5, 2, 3, 4, 5, or more. In another embodiment, the hazard ratios for the prognosis genes thus identified are no more than 0.67, 0.5., 0.33, 0.25., 0.2, or less.

[0075] Other rank tests, scores, measurements, or models can also be employed to identify prognosis genes whose expression profiles in peripheral blood samples are correlated with clinical outcome of solid tumors. These tests, scores, measurements, or models can be either parametric or nonparametric, and the regression may be either linear or non-linear. Many statistical methods and correlation/regression models can be carried out using commercially available programs.

[0076] Other methods capable of identifying genes differentially expressed in peripheral blood cells of one class of patients relative to another class of patients can be used. These methods include, but are not limited, RT-PCR, Northern Blot, *in situ* hybridization, and immunoassays such as ELISA, RIA or Western Blot. The expression levels of genes thus identified can be substantially higher or substantially lower in peripheral blood cells (e.g., PBMCs) of one class of patients than in another class of patients. In some cases, the average peripheral blood expression level of a prognosis gene

in PBMCs of one class of patients can be at least 2, 3, 4, 5, 10, 20, or more folds higher or lower than that in another class of patients. In many embodiments, the p-value of an appropriate statistical significance test (e.g., Student's t-test) for the difference between average expression levels is no more than 0.05, 0.01, 0.005, 0.001, 0.0005, 0.0001, or less.

[0077] Prognosis genes for other non-blood diseases can be similarly identified according to the present invention, provided that the correlation between peripheral blood gene expression and clinical outcome of these diseases is statistically significant. The peripheral blood expression patterns of the prognosis genes thus identified are indicative of clinical outcome of these diseases.

II. Identification of RCC Prognosis Genes

[0078] RCC comprises the majority of all cases of kidney cancer and is one of the ten most common cancers in industrialized countries, comprising 2% of adult malignancies and 2% of cancer-related deaths. Several prognostic factors and scoring indices have been developed for patients diagnosed with RCC, typified by multivariate assessments of several key indicators. As an example, one prognostic scoring system employs the five prognostic factors proposed by Motzer, *et al.*, *supra*– namely, Karnofsky performance status, serum lactate dehydrognease, hemoglobin, serum calcium, and presence/absence of prior nephrectomy.

[0079] The present invention identifies numerous RCC prognosis genes whose peripheral blood expression profiles correlate with patient outcome in CCI-779 therapy. In a clinical trial, the cytostatic mTOR inhibitor CCI-779 was evaluated in RCC patients for its anti-cancer effect. PBMCs collected prior to CCI-779 therapy were analyzed on oligonucleotide arrays in order to determine whether mononuclear cells from RCC patients possessed transcriptional patterns predictive of patient outcome. The results of both supervised and unsupervised analyses indicated that transcriptional profiles in the surrogate tissue of PBMCs from RCC patients prior to treatment with CCI-779 are significantly correlated with patient outcome.

[0080] PBMCs were isolated prior to CCI-779 therapy from peripheral blood of 45 advanced RCC patients (18 females and 27 males) participating in a phase 2 clinical trial study. Written informed consent for the pharmacogenomic portion of the clinical study was received for all individuals and the project was approved by the local Institutional Review

Boards at the participating clinical sites. RCC tumors of patients were classified at the clinical sites as conventional (clear cell) carcinomas (24), granular (1), papillary (3), or mixed subtypes (7). Ten tumors were classified as unknown. RCC patients were primarily of Caucasian descent (44 Caucasian, 1 African-American) and had a mean age of 58 years (range of 40 - 78 years). Inclusion criteria included patients with histologically confirmed advanced renal cancer who had received prior therapy for advanced disease, or who had not received prior therapy for advanced disease but were not appropriate candidates to receive high doses of IL-2 therapy. Other inclusion criteria included patients with (1) bi-dimensionally measurable evidence of disease; (2) evidence of progression of the disease prior to study entry; (3) an age of 18 years or older; (4) ANC > 1500/μL, platelet > 100,000/μL and hemoglobin > 8.5 g/dL; (5) adequate renal function evidenced by serum creatinine < 1.5 x upper limit of normal; (6) adequate hepatic function evidenced by biliruubin < 1.5 x upper limit of normal and AST < 3x upper limit of normal (or AST < 5x upper limit of normal if liver metastases were present); (7) serum cholesterol < 350 mg/dL, triglycerides < 300 mg/dL; (8) ECOG performance status 0-1; and (9) a life expectancy of at least 12 weeks. Exclusion criteria included patients who had (1) the presence of known CNS metastases; (2) surgery or radiotherapy within 3 weeks of start of dosing; (3) chemotherapy or biologic therapy for RCC within 4 weeks of start of dosing; (4) treatment with a prior investigational agent within 4 weeks of start of dosing; (5) immunocompromised status including those known to be HIV positive, or receiving concurrent use of immunosuppressive agents including corticosteroids; (6) active infections; (7) required treatment with anticonvulsant therapy; (8) presence of unstable angina/myocardial infarction within 6 months/ongoing treatment of life-threatening arrythmia; (9) history of prior malignancy in past 3 years; (10) hypersensitivity to macrolide antibiotics; and (11) pregnancy or any other illness which would substantially increase the risk associated with participation in the study.

[0081] These advanced RCC patients were treated with one of 3 doses of CCI-779 (25 mg, 75 mg, or 250 mg) administered as a 30 minute intravenous (IV) infusion once weekly for the duration of the trial. CCI-779 is an ester analog of the immunosuppressant rapamycin and as such is a potent, selective inhibitor of the mammalian target of rapamycin. The mammalian target of rapamycin (mTOR) activates multiple signaling pathways, including phosphorylation of p70s6kinase, which results in increased translation of 5' TOP mRNAs encoding proteins involved in translation and entry into the G1 phase of the cell

cycle. By virtue of its inhibitory effects on mTOR and cell cycle control, CCI-779 functions as a cytostatic and immunosuppressive agent.

[0082] Clinical staging and size of residual, recurrent or metastatic disease were recorded prior to treatment and every 8 weeks following initiation of CCI-779 therapy. Tumor size was measured in centimeters and reported as the product of the longest diameter and its perpendicular. Measurable disease was defined as any bidimensionally measurable lesion where both diameters > 1.0 cm by CT-scan, X-ray or palpation. Tumor response was determined by the sum of the products of all measurable lesions. The categories for assignment of clinical response were given by the clinical protocol definitions (i.e., progressive disease, stable disease, minor response, partial response, and complete response). The category for assignment of prognosis under the Motzer risk assessment (favorable vs intermediate vs poor) was also used. Among the 45 RCC patients, 6 were assigned a favorable risk assessment, 17 patients possessed an intermediate risk score, and 22 patients received a poor prognosis classification. In addition to the categorical classifications, overall survival and time to disease progression were also monitored as clinical endpoints.

[0083] HgU95A genechips (manufactured by Affymetrix) were used to detect baseline expression profiles in PBMCs of the RCC patients prior to the CCI-779 therapy. Each HgU95A genechip comprises over 12,600 human sequences according to the Affymetrix Expression Analysis Technical Manual. RNA transcripts were first isolated from PBMCs of the RCC patients. cRNA was then prepared and hybridized to the genechips according to protocols described in the Affymetrix's Expression Analysis Technical Manual. Hybridization signals were collected, scaled, and normalized before being subject to further analysis. In one example, the log of the expression level for each gene was normalized across the samples such that the mean is zero and the standard deviation is one.

[0084] The expression profiling analysis revealed that of the 12,626 genes on the HgU95A chip, 5,424 genes met the initial criteria (i.e., at least 1 present call across the data set and at least 1 frequency $\geq$ 10 ppm). On average, 4,023 transcripts were detected as "present" in any given RCC PBMC profile.

[0085] In an initial assessment of the expression data in baseline PBMCs, pairwise correlations were calculated to assess the association between gene expression levels measured by HgU95A Affymetrix microarrays and continuous measures of clinical

outcome. Correlations were run using expression levels from each of 5,424 qualifiers that passed the initial criteria. Correlations were run for two clinical measures (TTD and TTP) and for one measure of baseline expression level (log_2-transformed scaled frequency in units of ppm).

[0086] In one example, Spearman's rank correlations were computed. The p-value for the hypothesis that the correlation was equal to 0 was calculated for each pairwise correlation. For each comparison between clinical outcome and gene expression, the number of tests that were nominally significant out of the 5,424 tests performed was calculated for five Type I (i.e. false-positive) error levels. To adjust for the fact that 5,424 non-independent tests were performed, a permutation-based approach was employed to evaluate how often the observed number of significance tests would be found under the null hypothesis of no correlation.

[0087] The overall results for Spearman's rank correlation comparisons of clinical outcome with baseline expression levels (log_2-transformed scaled frequency) are summarized in Tables 2a and 2b. Each table shows alpha confidence levels ("α"), the observed numbers of transcripts that have nominally significant Spearman correlations with the clinical outcome of interest ("Observed Number"), and the percentage of permutations for which number of nominally significant Spearman correlations equals or exceeds the number observed ("%-age of Permutations"). Evidence for association between clinical outcome and baseline gene expression in PBMCs was significant for both TTD and TTP.

Table 2a. Spearman Correlations of Clinical Outcome with Baseline Expression Levels in PBMCs of RCC Patients in CCI-779 Therapy (n = 45 patients)

Time to Disease Progression		
α	Observed Number of Nominally Significant Spearman Correlations*	%-age of Permutations for which Number of Nominally Significant Spearman Correlations equals or exceeds observed number
0.1	1127	5.3% (53/1000)
0.05	749	3.8% (38/1000)
0.01	248	3.1% (31/1000)
0.005	159	2.6% (26/1000)
0.001	51	2.5% (25/1000)

* based on 5,424 genes (filtered by at least one Present and at least one frequency ≥ 10 ppm)

Table 2b. Spearman Correlations of Clinical Outcome with Baseline Expression Levels in PBMCs of RCC Patients in CCI-779 Therapy (n = 45 patients)

Time to Death		
α	Observed Number of Nominally Significant Spearman Correlations*	%-age of Permutations for which Number of Nominally Significant Spearman Correlations equals or exceeds observed number
0.1	1604	0.1% (1/1000)
0.05	1117	0.1% (1/1000)
0.01	436	0.1% (1/1000)
0.005	289	0.1% (1/1000)
0.001	105	0.3% (3/1000)

* based on 5,424 genes (filtered by at least one Present and at least one frequency ≥ 10 ppm)

[0088] Table 3 lists the results of the Spearman's rank correlation analyses for all of the 5,424 genes that met the initial criteria. Each gene has a corresponding qualifier on the HgU95A genechip, and each qualifier represents multiple oligonucleotide probes that are stably attached to discrete regions on the HgU95A genechip. According to the design, RNA transcripts of a gene, or the complements thereof, are expected to hybridize under nucleic acid array hybridization conditions to the corresponding qualifier on the HgU95A genechip. As used herein, a polynucleotide can hybridize to a qualifier if the polynucleotide, or the complement thereof, can hybridize to at least one oligonucleotide probe of the qualifier. In many embodiments, the polynucleotide or the complement thereof can hybridize to at least 50%, 60%, 70%, 80%, 90% or 100% of all of the oligonucleotide probes of the qualifier.

[0089] Each gene or qualifier in Table 3 may have a corresponding SEQ ID NO or Entrez accession number from which the oligonucleotide probes of the qualifier can be derived. In many instances, a polypeptide capable of hybridizing to a qualifier can also hybridize to the sequence of the corresponding SEQ ID NO or Entrez accession number, or the complement thereof. The sequence of each Entrez accession number can be obtained from the Entrez nucleotide database at the National Center of Biotechnology Information (NCBI). The Entrez nucleotide database collects sequences from several sources, including GenBank, RefSeq, and PDB. Each SEQ ID NO may be derived from the sequence of the corresponding Entrez accession number. Table 4 shows the Entrez and Unigene accession numbers for all of the qualifiers on the HgU95A genechip that met the initial criteria.

[0090] Any ambiguous residue ("n") in a SEQ ID NO can be determined by a variety of methods. In one embodiment, the ambiguous residues in a SEQ ID NO are determined by aligning the SEQ ID NO to a corresponding genomic sequence obtained from a human genome sequence database. In another embodiment, the ambiguous residues in a SEQ ID NO are determined based on the sequence of the corresponding Entrez accession number. In yet another embodiment, the ambiguous residues are determined by re-sequencing the SEQ ID NO.

[0091] Genes associated with each qualifier on the HgU95A genechip can be identified based on the annotations provided by Affymetrix. All of the genes thus identified are listed in Tables 3 and 5. These genes can also be identified based on their corresponding Entrez or Unigene accession numbers. In addition, these genes can be determined by BLAST searching their corresponding SEQ ID NOs, or the unambiguous segments thereof, against a human genome sequence database. Suitable human genome sequence databases for this purpose include, but are not limited to, the NCBI human genome database. The NCBI provides BLAST programs, such as "blastn," for searching its sequence databases.

[0092] In one embodiment, the BLAST search of the NCBI human genome database is carried out by using an unambiguous segment (e.g., the longest unambiguous segment) of a SEQ ID NO. Gene(s) that aligns to the unambiguous segment with significant sequence identity can be identified. In many cases, the identified gene(s) has at least 95%, 96%, 97%, 98%, 99%, or more sequence identity with the unambiguous segment.

[0093] On the basis of Spearman's rank correlation, prognosis genes that are highly correlated with TTP or TTD were identified. Table 6a lists examples of genes whose expression levels are positively correlated with TTP. Table 6b depicts examples of genes whose expression levels are negatively correlated with TTP. Table 6c provides examples of genes whose expression levels are positively correlated with TTD. Table 6d shows examples of genes whose expression levels are negatively correlated with TTD. Correlation coefficients, p-values, and the corresponding qualifiers are also indicated for each gene in Tables 6a, 6b, 6c, and 6d.

Table 6a. Prognosis Genes Positively Correlated with TTP

HgU95A Qualifier	Correlation Coefficient	P-Value	Gene Name
38518_at	0.6019	0.0000	SCML2
37343_at	0.5932	0.0000	ITPR3
41174_at	0.5925	0.0000	RANBP2L1
41669_at	0.5908	0.0000	KIAA0191
40584_at	0.5602	0.0001	NUP88
41767_r_at	0.5591	0.0001	KIAA0855
38256_s_at	0.5551	0.0001	DKFZP564O092
39829_at	0.5508	0.0001	ARL7
35802_at	0.5475	0.0001	KIAA1014
32169_at	0.5407	0.0001	KIAA0875
41562_at	0.5272	0.0002	BMI1
35753_at	0.5226	0.0002	PRP8
40905_s_at	0.5223	0.0002	DKFZP566J153
41547_at	0.5189	0.0003	BUB3
37416_at	0.5177	0.0003	ARHH
37585_at	0.5157	0.0003	SNRPA1
34716_at	0.5143	0.0003	TASR
32183_at	0.5034	0.0004	SFRS11
39426_at	0.4977	0.0005	CA150
35815_at	0.4975	0.0005	HYPB
36403_s_at	0.4972	0.0005	UNK_AI434146
40828_at	0.4963	0.0005	P85SPR
35364_at	0.4947	0.0006	APPBP1
33861_at	0.4931	0.0006	UNK_AI123426
36474_at	0.4927	0.0006	KIAA0776
35764_at	0.4908	0.0006	CXORF5
39129_at	0.4904	0.0006	UNK_AF052134
32508_at	0.4893	0.0006	KIAA1096
35842_at	0.4862	0.0007	UNK_AL049265
41737_at	0.4862	0.0007	SRM160
36303_f_at	0.4833	0.0008	ZNF85
34256_at	0.4829	0.0008	SIAT9
33845_at	0.4828	0.0008	HNRPH1
40048_at	0.4822	0.0008	UNK_D43951

HgU95A Qualifier	Correlation Coefficient	P-Value	Gene Name
37625_at	0.4801	0.0008	IRF4
33234_at	0.4779	0.0009	UNK_AA887480
2000_at	0.4777	0.0009	ATM
37078_at	0.4760	0.0010	CD3Z
38778_at	0.4744	0.0010	KIAA1046

Table 6b. Prognosis Genes Negatively Correlated with TTP

HgU95A Qualifier	Correlation Coefficient	P-Value	Gene Name
935_at	-0.6319	0.0000	CAP
34498_at	-0.5385	0.0001	VNN2
37023_at	-0.5292	0.0002	LCP1
286_at	-0.5189	0.0003	H2AFO
38831_f_at	-0.5152	0.0003	UNK_AF053356
268_at	-0.5126	0.0003	PECAM1
38893_at	-0.5006	0.0005	NCF4
34319_at	-0.4950	0.0005	S100P
37328_at	-0.4931	0.0006	PLEK
181_g_at	-0.4925	0.0006	UNK_S82470
38894_g_at	-0.4852	0.0007	NCF4
32736_at	-0.4805	0.0008	UNK_W68830

Table 6c. Prognosis Genes Positively Correlated with TTD

HgU95A Qualifier	Correlation Coefficient	P-Value	Gene Name
37385_at	0.6524	0.0000	CYP
41606_at	0.6155	0.0000	DRG1
33420_g_at	0.6043	0.0000	API5
35353_at	0.5969	0.0000	PSMC2
38017_at	0.5942	0.0000	CD79A
31851_at	0.5854	0.0000	RFP2
35319_at	0.5817	0.0000	CTCF
38702_at	0.5702	0.0000	UNK_AF070640
36474_at	0.5654	0.0001	KIAA0776
34256_at	0.5649	0.0001	SIAT9
34763_at	0.5575	0.0001	CSPG6
33831_at	0.5561	0.0001	CREBBP

HgU95A Qualifier	Correlation Coefficient	P-Value	Gene Name
229_at	0.5499	0.0001	CBF2
37381_g_at	0.5478	0.0001	GTF2B
40092_at	0.5436	0.0001	BAZ2A
39746_at	0.5428	0.0001	POLR2B
41174_at	0.5424	0.0001	RANBP2L1
32508_at	0.5397	0.0001	KIAA1096
33403_at	0.5390	0.0001	DKFZP547E1010
39809_at	0.5381	0.0001	HBP1
34829_at	0.5373	0.0001	DKC1
37625_at	0.5350	0.0002	IRF4
35656_at	0.5336	0.0002	RNF6
39509_at	0.5328	0.0002	UNK_AI692348
33543_s_at	0.5324	0.0002	PNN
38082_at	0.5318	0.0002	KIAA0650
36303_f_at	0.5311	0.0002	ZNF85
1885_at	0.5300	0.0002	ERCC3
32194_at	0.5285	0.0002	CBF2
41621_i_at	0.5264	0.0002	ZNF266
33151_s_at	0.5239	0.0002	UNK_W25932
32169_at	0.5212	0.0002	KIAA0875
36845_at	0.5203	0.0002	KIAA0136
36231_at	0.5197	0.0003	UNK_AC002073
35163_at	0.5172	0.0003	KIAA1041
40905_s_at	0.5170	0.0003	DKFZP566J153
39431_at	0.5164	0.0003	NPEPPS
41669_at	0.5160	0.0003	KIAA0191
35294_at	0.5150	0.0003	SSA2
39401_at	0.5139	0.0003	UNK_W28264
34716_at	0.5137	0.0003	TASR
40563_at	0.5136	0.0003	DKFZP564A043
38667_at	0.5124	0.0003	UNK_AA189161
38122_at	0.5107	0.0003	SLC23A1
37585_at	0.5096	0.0004	SNRPA1
32183_at	0.5079	0.0004	SFRS11
40816_at	0.5074	0.0004	PWP1

HgU95A Qualifier	Correlation Coefficient	P-Value	Gene Name
33818_at	0.5055	0.0004	UNK_AC004472
37703_at	0.5042	0.0004	RABGGTB
38016_at	0.5039	0.0004	HNRPD
37737_at	0.4997	0.0005	PCMT1
36872_at	0.4976	0.0005	ARPP-19
39415_at	0.4975	0.0005	HNRPK
40252_g_at	0.4970	0.0005	HRB2
39727_at	0.4966	0.0005	DUSP11
1728_at	0.4966	0.0005	BMI1
34967_at	0.4956	0.0005	UNK_AF001549
39864_at	0.4949	0.0005	CIRBP
32758_g_at	0.4947	0.0006	RAE1
35753_at	0.4943	0.0006	PRP8
1857_at	0.4916	0.0006	MADH7
35764_at	0.4915	0.0006	CXORF5
32372_at	0.4911	0.0006	CTSB
33485_at	0.4892	0.0006	RPL4
34647_at	0.4887	0.0007	DDX5
1442_at	0.4886	0.0007	ESR2
41506_at	0.4875	0.0007	MAPKAPK5
34879_at	0.4873	0.0007	DPM1
39512_s_at	0.4869	0.0007	UNK_AA457029
36783_f_at	0.4865	0.0007	H-PLK
35479_at	0.4860	0.0007	ADAM28
40308_at	0.4858	0.0007	UNK_AI830496
38462_at	0.4852	0.0007	NDUFA5
781_at	0.4851	0.0007	RABGGTB
38102_at	0.4850	0.0007	UNK_W28575
38256_s_at	0.4829	0.0008	DKFZP564O092
32850_at	0.4817	0.0008	NUP153
35286_r_at	0.4815	0.0008	RY1
36456_at	0.4815	0.0008	DKFZP564I052
38924_s_at	0.4813	0.0008	SSH3BP1
35805_at	0.4809	0.0008	DKFZP434D156
40086_at	0.4805	0.0008	KIAA0261

HgU95A Qualifier	Correlation Coefficient	P-Value	Gene Name
34274_at	0.4801	0.0008	KIAA1116
39897_at	0.4793	0.0009	DDX16
41665_at	0.4792	0.0009	KIAA0824
38114_at	0.4785	0.0009	RAD21
41166_at	0.4782	0.0009	IGHM
41569_at	0.4781	0.0009	KIAA0974
33440_at	0.4774	0.0009	TCF8
36459_at	0.4767	0.0009	KIAA0879
216_at	0.4765	0.0009	PTGDS
41199_s_at	0.4760	0.0009	SFPQ
40051_at	0.4756	0.0010	KIAA0057
38019_at	0.4754	0.0010	CSNK1E
36690_at	0.4746	0.0010	NR3C1
41547_at	0.4742	0.0010	BUB3
38105_at	0.4734	0.0010	UNK_W26521
40828_at	0.4732	0.0010	P85SPR
41809_at	0.4729	0.0010	UNK_AI656421
36210_g_at	0.4727	0.0010	FSRG1

Table 6d. Prognosis Genes Negatively Correlated with TTD

HgU95A Qualifier	Correlation Coefficient	P-Value	Gene Name
286_at	-0.5871	0.0000	H2AFO
32609_at	-0.5841	0.0000	H2AFO
38483_at	-0.5464	0.0001	HSA011916
769_s_at	-0.5036	0.0004	ANXA2
1131_at	-0.4876	0.0007	MAP2K2
32378_at	-0.4818	0.0008	PKM2
956_at	-0.4770	0.0009	TUBB
37311_at	-0.4760	0.0010	TALDO1
37148_at	-0.4744	0.0010	LILRB3
36199_at	-0.4725	0.0010	DAP

[0094] In addition to the specific genes described herein, the present invention contemplates the use of any other gene that can hybridize under stringent or nucleic acid array hybridization conditions to a qualifier identified in the present invention. These genes

may include hypothetical or putative genes that are supported by EST or mRNA data. The expression profiles of these genes may correlate with patient clinical outcome. As used herein, a gene can hybridize to a qualifier if an RNA transcript of the gene can hybridize to at least one oligonucleotide probe of the qualifier. In many cases, an RNA transcript of the gene can hybridize to at least 50%, 60%, 70%, 80%, 90%, or more oligonucleotide probes of the qualifier.

[0095] The oligonucleotide probe sequences of each qualifier on HgU95A genechips may be obtained from Affymetrix or from the sequence files maintained at Affymetrix website "www.affymetrix.com/support/technical/byproduct.affx?product= hgu95sequence." For instance, the oligonucleotide probe sequences can be found in the sequence file "HG_U95A Probe Sequences, FASTA" at the website. This sequence file is incorporated herein by reference in its entirety.

[0096] In another example, a Cox proportional hazard regression model was employed to assess the correlation between baseline PBMC gene expression levels and clinical outcome. Cox model can take into account the effects of censoring on correlations of gene expression with TTD (or Survival as of last known date alive) and TTP (or progression-free status as of last known date alive). Of the 45 RCC patients with baseline PBMC expression levels, 4 had censored data for TTP and 15 had censored data for TTD. Similar to the Spearman's assessment of the data, Cox regression can identify genes significantly correlated with survival and disease progression for any given α-confidence level. A similar permutation strategy can be used to affirm any correlation between baseline expression profiles and clinical outcome.

[0097] In one embodiment, models were fit using expression levels from each of the 5,424 qualifiers that passed the initial filtering criteria in the 45 baseline samples. TTP and TTD were tested for their association with log2-transformed scaled frequency at baseline. A SAS program was used to generate the estimates in Tables 7a and 7b. Tables 7a and 7b demonstrate a strong correlation between TTP/TTD and baseline gene expression.

Table 7a. Cox Regressions of Clinical Outcome on Baseline Expression Levels in PBMCs of RCC Patients in CCI-779 Therapy (n = 45 patients)

Time to Progression		
∀	Observed Number of Nominally Significant	Percentage of Permutations for which Number of Nominally

	Cox Regressions*	Significant Cox Regressions Equals or Exceeds Observed Number**
0.1	1439	0.8% (4/500)
0.05	950	0.8% (3/500)
0.01	342	0.8% (4/500)
0.005	217	0.8% (4/500)
0.001	53	1.0% (5/500)

* for 5,424 genes (filtered by at least one Present call and at least one frequency ≥ 10 ppm)
** based on 500 random permutations

Table 7b. Cox Regressions of Clinical Outcome on Baseline Expression Levels in PBMCs of RCC Patients in CCI-779 Therapy (n = 45 patients)

Time to Death		
∀	Observed Number of Nominally Significant Cox Regressions*	Percentage of Permutations for which Number of Nominally Significant Cox Regressions Equals or Exceeds Observed Number**
0.1	1948	<0.2% (0/500)
0.05	1383	<0.2% (0/500)
0.01	602	<0.2% (0/500)
0.005	404	<0.2% (0/500)
0.001	142	<0.2% (0/500)

* for 5,424 genes (filtered by at least one Present call and at least one frequency ≥ 10 ppm)
** based on 500 random permutations

[0098] Table 8 lists the results of Cox proportional hazard modeling for all of the 5,424 genes that met the initial criteria. Hazard ratios and p-values (for the hypothesis that the risk coefficient was equal to 1, i.e., no risk) are indicated for each gene. Examples of genes that are indicative of high risk for TTP or TTD are shown in Tables 9a or 9c, respectively. These genes have hazard ratios of at least 3. Examples of genes that are indicative of low risk for TTP or TTD are described in Tables 9b or 9d, respectively. These genes have hazard ratios of no more than 0.333.

Table 9a. Prognosis Genes Indicative of High Risk for TTP

HgU95A Qualifier	Hazard Ratio	P-Value	Gene Name
37023_at	6.1066	0.0001	LCP1
935_at	5.8829	0.0000	CAP
40771_at	4.9503	0.0586	MSN
37298_at	4.6595	0.0046	GABARAP

HgU95A Qualifier	Hazard Ratio	P-Value	Gene Name
31820_at	4.2099	0.0061	HCLS1
676_g_at	4.1051	0.0016	IFITM1
33906_at	3.9750	0.0106	SSSCA1
32736_at	3.8093	0.0013	UNK_W68830
40169_at	3.5692	0.0243	TIP47
39811_at	3.4197	0.1074	UNK_AA402538
1309_at	3.3680	0.0053	PSMB3
39814_s_at	3.2703	0.0029	UNK_AI052724
38605_at	3.1625	0.0592	NDUFB1
38831_f_at	3.0853	0.0092	UNK_AF053356

Table 9b. Prognosis Genes Indicative of Low Risk for TTP

HgU95A Qualifier	Hazard Ratio	P-Value	Gene Name
39415_at	0.0818	0.0002	HNRPK
35753_at	0.1608	0.0001	PRP8
33667_at	0.1650	0.0890	PPIA
33845_at	0.1657	0.0024	HNRPH1
36186_at	0.1661	0.0040	RNPS1
1420_s_at	0.1662	0.0009	EIF4A2
31950_at	0.1724	0.0071	PABPC1
34647_at	0.1831	0.0010	DDX5
36515_at	0.2094	0.0002	GNE
36111_s_at	0.2147	0.0031	SFRS2
39180_at	0.2154	0.0009	FUS
32758_g_at	0.2186	0.0010	RAE1
31952_at	0.2211	0.0076	RPL6
38527_at	0.2258	0.0016	NONO
32831_at	0.2298	0.0006	TIM17
37609_at	0.2321	0.0016	NUBP1
34695_at	0.2330	0.0035	GA17
39730_at	0.2331	0.0005	ABL1
35808_at	0.2385	0.0037	SFRS6
32751_at	0.2386	0.0013	UNK_AF007140
41737_at	0.2393	0.0023	SRM160
32205_at	0.2431	0.0009	PRKRA

HgU95A Qualifier	Hazard Ratio	P-Value	Gene Name
40252_g_at	0.2473	0.0033	HRB2
35325_at	0.2540	0.0030	UNK_AF052113
41292_at	0.2549	0.0014	HNRPH1
32658_at	0.2553	0.0010	UNK_AL031228
33307_at	0.2569	0.0008	UNK_AL022316
40426_at	0.2587	0.0306	BCL7B
41562_at	0.2595	0.0010	BMI1
34315_at	0.2638	0.0149	AFG3L2
33920_at	0.2665	0.0549	DIAPH1
33706_at	0.2698	0.0114	SART1
35170_at	0.2706	0.0053	MAN2C1
229_at	0.2715	0.0064	CBF2
33485_at	0.2724	0.0169	RPL4
1728_at	0.2736	0.0103	BMI1
38105_at	0.2748	0.0017	UNK_W26521
1361_at	0.2801	0.0059	TERF1
32171_at	0.2831	0.0040	EIF5
36456_at	0.2834	0.0015	DKFZP564I052
838_s_at	0.2841	0.0616	UBE2I
1706_at	0.2852	0.0144	ARAF1
38778_at	0.2882	0.0012	KIAA1046
39378_at	0.2896	0.1463	BECN1
34225_at	0.2911	0.0126	UNK_AF101434
32833_at	0.2918	0.0016	CLK1
34285_at	0.2938	0.0021	KIAA0795
35743_at	0.2968	0.0133	NAR
39165_at	0.2971	0.0086	NIFU
36685_at	0.2979	0.0045	AMD1
37557_at	0.2985	0.0038	SLC4A2
36303_f_at	0.2987	0.0018	ZNF85
33392_at	0.3019	0.0030	DKFZP434J154
40160_at	0.3031	0.0038	DKFZP586P2220
34337_s_at	0.3047	0.0009	M96
37506_at	0.3053	0.0006	UNK_Z78308
38256_s_at	0.3053	0.0002	DKFZP564O092

HgU95A Qualifier	Hazard Ratio	P-Value	Gene Name
37690_at	0.3053	0.0120	ILVBL
1020_s_at	0.3060	0.0069	SIP2-28
36862_at	0.3066	0.0147	KIAA1115
39141_at	0.3069	0.0074	ABCF1
32592_at	0.3071	0.0280	KIAA0323
39044_s_at	0.3076	0.0141	DGKD
40596_at	0.3076	0.0058	TCOF1
34369_at	0.3078	0.0454	KIAA0214
33188_at	0.3090	0.0006	PPIL2
41220_at	0.3110	0.0404	MSF
38445_at	0.3125	0.0057	ARHGEF1
36783_f_at	0.3125	0.0064	H-PLK
37717_at	0.3126	0.0130	NAGR1
36198_at	0.3167	0.0058	KIAA0016
35125_at	0.3171	0.0540	RPS6
32438_at	0.3172	0.0557	RPS20
37030_at	0.3181	0.0006	KIAA0887
37703_at	0.3183	0.0011	RABGGTB
1711_at	0.3199	0.0463	TP53BP1
41691_at	0.3216	0.0006	KIAA0794
32079_at	0.3219	0.0037	KIAA0639
39865_at	0.3230	0.0151	UNK_AI890903
34326_at	0.3232	0.0025	COPB
34808_at	0.3244	0.0188	KIAA0999
36129_at	0.3244	0.0014	UNK_AB007857
37672_at	0.3249	0.0077	USP7
32208_at	0.3257	0.0098	KIAA0355
35298_at	0.3266	0.0973	EIF3S7
36982_at	0.3267	0.0018	USP14
31573_at	0.3292	0.0566	RPS25
36603_at	0.3292	0.0015	GCN1L1
36189_at	0.3310	0.0661	ILF2
39155_at	0.3325	0.0433	PSMD3

Table 9c. Prognosis Genes Indicative of High Risk for TTD

HgU95A Qualifier	Hazard Ratio	P-Value	Gene Name
40771_at	9.6763	0.0122	MSN
39811_at	8.0370	0.0149	UNK_AA402538
37298_at	7.6453	0.0021	GABARAP
38483_at	6.7764	0.0001	HSA011916
1878_g_at	6.1122	0.0004	ERCC1
33994_g_at	4.9451	0.0009	MYL6
32318_s_at	4.9169	0.0027	ACTB
37012_at	4.8396	0.0057	CAPZB
1199_at	4.7016	0.0103	EIF4A1
36641_at	4.5981	0.0042	CAPZA2
34160_at	4.5693	0.0086	ACTG1
34091_s_at	4.4114	0.0158	VIM
286_at	4.2492	0.0000	H2AFO
35770_at	4.1617	0.0083	ATP6S1
33341_at	4.0632	0.0102	GNB1
33659_at	4.0505	0.0074	CFL1
935_at	4.0159	0.0016	CAP
40134_at	3.8316	0.0043	ATP5J2
37346_at	3.8205	0.0126	ARF5
37023_at	3.8170	0.0059	LCP1
38451_at	3.8077	0.0034	UQCR
34836_at	3.7786	0.0080	RABL
35263_at	3.6729	0.0558	EIF4EBP2
41724_at	3.6595	0.0026	DXS1357E
33679_f_at	3.5643	0.0134	TUBB2
33121_g_at	3.5151	0.0007	RGS10
40872_at	3.4884	0.0013	COX6B
1315_at	3.4428	0.0026	UNK_D78361
36574_at	3.4083	0.1032	IDH3G
1131_at	3.3872	0.0002	MAP2K2
31444_s_at	3.3199	0.0016	ANXA2P2
36963_at	3.3124	0.0060	PGD
35083_at	3.2546	0.0517	UNK_AL031670
32145_at	3.2308	0.0012	ADD1
AFFX-HSAC07/X00351_3_at	3.1377	0.0060	BACTIN3_Hs_AFFX

HgU95A Qualifier	Hazard Ratio	P-Value	Gene Name
769_s_at	3.1358	0.0006	ANXA2
35783_at	3.0738	0.0592	UNK_H93123
32609_at	3.0361	0.0000	H2AFO
1695_at	3.0329	0.0225	NEDD8

Table 9d. Prognosis Genes Indicative of Low Risk for TTD

HgU95A Qualifier	Hazard Ratio	P-Value	Gene Name
41606_at	0.0322	0.0000	DRG1
38016_at	0.0547	0.0003	HNRPD
39274_at	0.1030	0.0004	NUP62
36189_at	0.1100	0.0029	ILF2
35353_at	0.1140	0.0000	PSMC2
1728_at	0.1250	0.0001	BMI1
40252_g_at	0.1265	0.0003	HRB2
36210_g_at	0.1287	0.0003	FSRG1
34315_at	0.1288	0.0028	AFG3L2
34647_at	0.1295	0.0001	DDX5
38702_at	0.1333	0.0000	UNK_AF070640
39415_at	0.1428	0.0019	HNRPK
33818_at	0.1433	0.0011	UNK_AC004472
37509_at	0.1447	0.0001	UNK_AF046059
31952_at	0.1466	0.0025	RPL6
37385_at	0.1538	0.0000	CYP
33485_at	0.1591	0.0010	RPL4
34695_at	0.1620	0.0013	GA17
37609_at	0.1625	0.0004	NUBP1
32807_at	0.1675	0.0012	DKFZP566C134
33614_at	0.1694	0.0017	RPL18A
32758_g_at	0.1727	0.0010	RAE1
32766_at	0.1742	0.0056	G22P1
36872_at	0.1763	0.0001	ARPP-19
34401_at	0.1764	0.0095	UQCRFS1
36186_at	0.1791	0.0047	RNPS1
35319_at	0.1792	0.0000	CTCF
755_at	0.1796	0.0023	ITPR1

HgU95A Qualifier	Hazard Ratio	P-Value	Gene Name
40370_f_at	0.1809	0.0104	HLA-G
37353_g_at	0.1824	0.0013	SP100
41295_at	0.1825	0.0005	GPX3
36845_at	0.1886	0.0001	KIAA0136
229_at	0.1887	0.0008	CBF2
39766_r_at	0.1906	0.0016	POLR2K
40426_at	0.1909	0.0183	BCL7B
38456_s_at	0.1912	0.0240	UNK_AL049650
35595_at	0.1945	0.0000	CGRP-RCP
35656_at	0.1945	0.0001	RNF6
35753_at	0.1955	0.0014	PRP8
37367_at	0.1965	0.0429	ATP6E
38590_r_at	0.1981	0.0171	PTMA
35125_at	0.2004	0.0120	RPS6
37381_g_at	0.2014	0.0003	GTF2B
36946_at	0.2024	0.0004	DYRK1A
38068_at	0.2027	0.0010	AMFR
32175_at	0.2049	0.0156	CDC10
31538_at	0.2057	0.0031	RPLP0
39727_at	0.2079	0.0003	DUSP11
36456_at	0.2120	0.0003	DKFZP564I052
37672_at	0.2121	0.0013	USP7
41288_at	0.2154	0.0060	CALM1
38114_at	0.2167	0.0036	RAD21
33543_s_at	0.2190	0.0002	PNN
35325_at	0.2193	0.0043	UNK_AF052113
39562_at	0.2197	0.0018	CGGBP1
37737_at	0.2226	0.0004	PCMT1
33740_at	0.2241	0.0061	UNK_AF023268
1361_at	0.2250	0.0030	TERF1
1020_s_at	0.2250	0.0020	SIP2-28
38102_at	0.2281	0.0001	UNK_W28575
35294_at	0.2308	0.0003	SSA2
40700_at	0.2309	0.0022	SP140
39020_at	0.2310	0.0067	SIVA

HgU95A Qualifier	Hazard Ratio	P-Value	Gene Name
1449_at	0.2311	0.0025	PSMA4
34821_at	0.2319	0.0007	DKFZP586D0623
36783_f_at	0.2319	0.0010	H-PLK
39740_g_at	0.2329	0.0085	NACA
39155_at	0.2333	0.0138	PSMD3
39864_at	0.2344	0.0002	CIRBP
39099_at	0.2361	0.0011	SEC23A
32208_at	0.2365	0.0036	KIAA0355
39027_at	0.2377	0.0174	COX4
39774_at	0.2390	0.0207	OXA1L
40449_at	0.2391	0.0006	RFC1
40369_f_at	0.2395	0.0154	UNK_AL022723
33151_s_at	0.2407	0.0002	UNK_W25932
37625_at	0.2410	0.0000	IRF4
35055_at	0.2415	0.0223	BTF3
33845_at	0.2416	0.0065	HNRPH1
33451_s_at	0.2418	0.0128	RPL22
38527_at	0.2425	0.0064	NONO
40563_at	0.2425	0.0001	DKFZP564A043
36975_at	0.2427	0.0037	UNK_W26659
38854_at	0.2445	0.0037	KIAA0635
35163_at	0.2485	0.0001	KIAA1041
38817_at	0.2492	0.0087	SPAG7
41787_at	0.2502	0.0004	KIAA0669
649_s_at	0.2504	0.0001	CXCR4
37715_at	0.2510	0.0002	SNW1
33403_at	0.2511	0.0000	DKFZP547E1010
34172_s_at	0.2512	0.0013	UNK_M99578
32576_at	0.2522	0.0151	EIF3S5
39378_at	0.2550	0.1231	BECN1
35286_r_at	0.2554	0.0009	RY1
37350_at	0.2559	0.0102	UNK_AL031177
38123_at	0.2559	0.0025	D123
41506_at	0.2559	0.0001	MAPKAPK5
40140_at	0.2559	0.0004	ZFP103

HgU95A Qualifier	Hazard Ratio	P-Value	Gene Name
38073_at	0.2561	0.0018	RNMT
31872_at	0.2563	0.0029	SSXT
34349_at	0.2564	0.0035	SEC63L
39792_at	0.2568	0.0002	HNRPR
35187_at	0.2578	0.0061	UNK_AL080216
1220_g_at	0.2578	0.0003	IRF2
33706_at	0.2584	0.0209	SART1
34809_at	0.2588	0.0102	KIAA0999
39342_at	0.2588	0.0499	MARS
40874_at	0.2593	0.0541	EDF1
40814_at	0.2597	0.0009	IDS
39809_at	0.2597	0.0000	HBP1
37226_at	0.2599	0.0014	BNIP1
34370_at	0.2604	0.0020	ARCN1
40651_s_at	0.2604	0.0010	CRHR1
40816_at	0.2607	0.0004	PWP1
35195_at	0.2613	0.0051	RPC
40110_at	0.2621	0.0108	IDH3B
33886_at	0.2625	0.0019	SSH3BP1
34879_at	0.2639	0.0015	DPM1
36968_s_at	0.2660	0.0019	OIP2
36303_f_at	0.2669	0.0006	ZNF85
40219_at	0.2670	0.0103	HIS1
38942_r_at	0.2670	0.0105	UNK_W28610
32487_s_at	0.2672	0.0061	KPNA4
36754_at	0.2675	0.0001	ADCYAP1
39739_at	0.2683	0.0496	MYH9
33443_at	0.2687	0.0004	UNK_Z99129
31950_at	0.2687	0.0321	PABPC1
39059_at	0.2689	0.0145	DHCR7
33831_at	0.2702	0.0001	CREBBP
35368_at	0.2703	0.0006	ZNF207
35227_at	0.2706	0.0057	RBBP8
41296_s_at	0.2713	0.0009	GPX3
40596_at	0.2717	0.0047	TCOF1

HgU95A Qualifier	Hazard Ratio	P-Value	Gene Name
35910_f_at	0.2720	0.0113	MMPL1
34018_at	0.2722	0.0014	COL19A1
36949_at	0.2722	0.0033	CSNK1D
33394_at	0.2730	0.0011	DDX19
34231_at	0.2734	0.0036	UNK_AF074606
32288_r_at	0.2738	0.0014	KLRC3
38903_at	0.2742	0.0007	GJB5
38040_at	0.2743	0.0093	SPF30
39126_at	0.2749	0.0043	UNK_AL080101
35321_at	0.2752	0.0034	TLK2
36546_r_at	0.2755	0.0142	UNK_AB011114
39746_at	0.2755	0.0000	POLR2B
41256_at	0.2762	0.0054	EEF1D
41789_r_at	0.2781	0.0012	KIAA0669
35630_at	0.2784	0.0025	LLGL2
40984_at	0.2789	0.0384	UNK_W28255
35199_at	0.2789	0.0035	KIAA0982
40308_at	0.2791	0.0003	UNK_AI830496
40803_at	0.2793	0.0014	UNK_AL050161
322_at	0.2801	0.0045	PIK3R3
1885_at	0.2804	0.0008	ERCC3
193_at	0.2814	0.0330	TAF2G
38668_at	0.2819	0.0141	KIAA0553
39730_at	0.2819	0.0088	ABL1
38256_s_at	0.2821	0.0009	DKFZP564O092
39290_f_at	0.2832	0.0013	DKFZP564M2423
34326_at	0.2833	0.0020	COPB
38923_at	0.2838	0.0075	FRG1
34225_at	0.2845	0.0092	UNK_AF101434
35258_f_at	0.2846	0.0023	SFRS2IP
31546_at	0.2847	0.0090	RPL18
37659_at	0.2855	0.0180	IMMT
37717_at	0.2861	0.0090	NAGR1
32592_at	0.2862	0.0215	KIAA0323
35978_at	0.2871	0.0215	UNK_AF009242

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
11 November 2004 (11.11.2004)

PCT

(10) International Publication Number
WO 2004/097052 A3

(51) International Patent Classification[7]: C12Q 1/68, C07K 16/18, G01N 33/53

(21) International Application Number: PCT/US2004/013587

(22) International Filing Date: 29 April 2004 (29.04.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/466,067 29 April 2003 (29.04.2003) US
60/538,246 23 January 2004 (23.01.2004) US

(71) Applicant *(for all designated States except US)*: **WYETH** [US/US]; 5 Giralda Farms, Madison, NJ 07940 (US).

(71) Applicant and
(72) Inventor: **STRAHS, Andrew** [US/US]; 30 McKinley Street, Maynard, MA 01754 (US).

(72) Inventor: **TREPICCHIO, William, L.**; 21 Abbott Ridge Drive, Andover, MA 01810 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **BURCZYNSKI, Michael, E.** [US/US]; 71 Franklin Avenue, Swampscott, MA 01907 (US). **TWINE, Natalie, C.** [US/US]; 379 Shirley Hill Road, Goffstown, NH 03045 (US). **SLONIM, Donna, K.** [US/US]; 799 Dale Street, North Andover, MA 01845 (US). **IMMERMANN, Fred** [US/US]; 194 Haverstraw Road, Suffern, NY 10901 (US). **DORNER, Andrew, J.** [US/US]; 20 Baskin Road, Lexington, MA 02421 (US).

(74) Agent: **VAN DYKE, Raymond**; Nixon Peabody LLP, 401 9th Street, N.W., Washington, DC 20004 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

(88) Date of publication of the international search report:
14 April 2005

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHODS FOR PROGNOSIS AND TREATMENT OF SOLID TUMORS

(57) Abstract: Solid tumor prognosis genes, and methods, systems and equipment of using these genes for the prognosis and treatment of solid tumors. Prognosis genes for a solid tumor can be identified by the present invention. The expression profiles of these genes in peripheral blood mononuclear cells (PBMCs) are correlated with clinical outcome of the solid tumor. The prognosis genes of the present invention can be used as surrogate markers for predicting clinical outcome of a solid tumor in a patient of interest. These genes can also be used to select a treatment which has a favorable prognosis for the solid tumor of the patient of interest.

WO 2004/097052 A3

INTERNATIONAL SEARCH REPORT

PCT/US2004/013587

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12Q1/68 C07K16/18 G01N33/53

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12Q G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS, EMBASE, WPI Data, PAJ, EMBL

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	BURCZYNSKI ET AL.: "Pharmacogenomic expression profiling of renal cell carcinoma in a phase II trial of CCI-779: identification of surrogate markers of disease and predictors of outcome in the compartment of peripheral blood" EUROPEAN JOURNAL OF CANCER, PERGAMON PRESS, OXFORD, GB, vol. 38, November 2002 (2002-11), page S51, XP002295167 ISSN: 0959-8049 abstract	1-20

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
4 November 2004	25. 01. 2005

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hagenmaier, S

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DIPAOLA R S ET AL: "Phase I clinical and pharmacologic study of 13-cis-retinoic acid, interferon alfa, and paclitaxel in patients with prostate cancer and other advanced malignancies." JOURNAL OF CLINICAL ONCOLOGY : OFFICIAL JOURNAL OF THE AMERICAN SOCIETY OF CLINICAL ONCOLOGY. JUL 1999, vol. 17, no. 7, July 1999 (1999-07), pages 2213-2218, XP002295169 ISSN: 0732-183X	1-5, 10-17
Y	the whole document	6-9
X	"Product Catalogue" January 2001 (2001-01), AFFYMETRIX , XP002301464 Human Genome U95A set. page 1 page 11	19,20
Y		1-18
X	SU ANDREW I ET AL: "Large-scale analysis of the human and mouse transcriptomes." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA. 2 APR 2002, vol. 99, no. 7, 2 April 2002 (2002-04-02), pages 4465-4470, XP002301463 ISSN: 0027-8424	19,20
Y	the whole document & DATABASE SOURCE 18 October 2004 (2004-10-18), http://genome-www5.stanford.edu/cgi-bin/source/expressionSearch?option=cluster&criteria=Hs.171558&organism=Hs abstract	1-18
Y	ELIT LAURIE: "CCI-779 Wyeth." CURRENT OPINION IN INVESTIGATIONAL DRUGS (LONDON, ENGLAND : 2000) AUG 2002, vol. 3, no. 8, August 2002 (2002-08), pages 1249-1253, XP008037562 ISSN: 1472-4472 page 1252, line 3 the whole document	1-20
Y	WO 02/40000 A (WYETH CORP) 23 May 2002 (2002-05-23) the whole document	1-20

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	PERALBA JOSEP-MARIA ET AL: "Pharmacodynamic evaluation of the rapamycin ester CCI-779" PROCEEDINGS OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH ANNUAL MEETING, vol. 43, March 2002 (2002-03), pages 1000-1001, XP001182924 & 93RD ANNUAL MEETING OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH; SAN FRANCISCO, CALIFORNIA, USA; APRIL 06-10, 2002 ISSN: 0197-016X the whole document	1-20
Y	RININGER J A ET AL: "Differential gene expression technologies for identifying surrogate markers of drug efficacy and toxicity" DRUG DISCOVERY TODAY 01 DEC 2000 UNITED KINGDOM, vol. 5, no. 12, 1 December 2000 (2000-12-01), pages 560-568, XP002249959 ISSN: 1359-6446 the whole document	1-20
Y	SCHULZE-KOOPS H ET AL: "Persistent reduction in IL-6 mRNA in peripheral blood mononuclear cells of patients with rheumatoid arthritis after treatment with a monoclonal antibody to CD54 (ICAM-1)." CLINICAL AND EXPERIMENTAL IMMUNOLOGY. NOV 1996, vol. 106, no. 2, November 1996 (1996-11), pages 190-196, XP002295168 ISSN: 0009-9104 the whole document	1-20
Y	WO 01/81916 A (CORNELL RES FOUNDATION; FERRAN CHRISTINE; VASCONCELLOS LAURO; SUTHANTH) 1 November 2001 (2001-11-01) claim 18 the whole document page 50	1-20

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GOLUB T R ET AL: "MOLECULAR CLASSIFICATION OF CANCER: CLASS DISCOVERY AND CLASS PREDICTION BY GENE EXPRESSION MONITORING" SCIENCE, AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE,, US, vol. 286, 15 October 1999 (1999-10-15), pages 531-537, XP002943419 ISSN: 0036-8075 cited in the application the whole document	1-20
A	MOTZER R J ET AL: "Survival and prognostic stratification of 670 patients with advanced renal cell carcinoma." JOURNAL OF CLINICAL ONCOLOGY : OFFICIAL JOURNAL OF THE AMERICAN SOCIETY OF CLINICAL ONCOLOGY. AUG 1999, vol. 17, no. 8, August 1999 (1999-08), pages 2530-2540, XP002301153 ISSN: 0732-183X cited in the application the whole document	
P,Y	PERALBA JOSEP MARIA ET AL: "Pharmacodynamic Evaluation of CCI-779, an Inhibitor of mTOR, in Cancer Patients." CLINICAL CANCER RESEARCH : AN OFFICIAL JOURNAL OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH. 1 AUG 2003, vol. 9, no. 8, 1 August 2003 (2003-08-01), pages 2887-2892, XP002295172 ISSN: 1078-0432 the whole document	1-20
E,L	WO 2004/048933 A (STOVER JENNIFER A ; DORNER ANDREW (US); SLONI DONNA K (US); TWINE NATA) 10 June 2004 (2004-06-10) page 16, line 1 - line 9 the whole document	1-5, 11-14, 16,17, 19,20
E,L	WO 2004/072265 A (TWINE NATALIE ; DORNER ANDREW J (US); WYETH CORP (US); BURCZYNSKI MICH) 26 August 2004 (2004-08-26) the whole document	1-20

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2004/013587

Box II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-16, 19, 20 (all completely); 17, 18 (partially)

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: Invention 1 (claims 1-16,19,20 (all completely); 17,18 (both partially)):

A method comprising comparing an expression profile of at least one gene in a peripheral blood sample of a patient to at least one reference expression profile of said at least one gene wherein the patient has a solid tumor, and expression levels of each of said at least one gene in PBMCs of patients who have the solid tumor correlate with clinical outcomes of said patients, and wherein in particular the at least one gene is SCML2, the solid tumor is RCC and the clinical outcomes are measured by patient response to a CCI-779 therapy; a system according to claim 19; a nucleic acid or protein according to claim 20.

Inventions 2-850 (claims 17,18 (both partially)):

A method comprising comparing an expression profile of at least one gene in a peripheral blood sample of a patient to at least one reference expression profile of said at least one gene, wherein the patient has a solid tumor, and expression levels of each of said at least one gene in PBMCs of patients who have the solid tumor correlate with clinical outcomes of said patients, wherein the at least one gene is ITPR3, the solid tumor is RCC and the clinical outcomes are measured by patient response to a CCI-779 therapy.
..ibidem for Inventions 3-850 relating to a different gene selected from tables 6a, 6b, 6c, 6d, 9a, 9b, 9c, 9d, 10, 11, 12, 13, 16, 20 and 21.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0240000 A	23-05-2002	AU 2731302 A	27-05-2002
		BR 0115323 A	17-02-2004
		CA 2429020 A1	23-05-2002
		EP 1335725 A2	20-08-2003
		HU 0400521 A2	28-06-2004
		JP 2004517065 T	10-06-2004
		NO 20032181 A	14-05-2003
		WO 0240000 A2	23-05-2002
		US 2002091137 A1	11-07-2002
		US 2003153593 A1	14-08-2003
WO 0181916 A	01-11-2001	US 2003104371 A1	05-06-2003
		AU 5716101 A	07-11-2001
		WO 0181916 A2	01-11-2001
		US 2002132235 A1	19-09-2002
WO 2004048933 A	10-06-2004	WO 2004048933 A2	10-06-2004
		US 2004110221 A1	10-06-2004
		US 2004175743 A1	09-09-2004
		US 2004235020 A1	25-11-2004
		WO 2004072265 A2	26-08-2004
		WO 2004094673 A2	04-11-2004
WO 2004072265 A	26-08-2004	US 2004175743 A1	09-09-2004
		WO 2004072265 A2	26-08-2004
		WO 2004048933 A2	10-06-2004
		US 2004110221 A1	10-06-2004
		US 2004235020 A1	25-11-2004
		WO 2004094673 A2	04-11-2004
		WO 2004097052 A2	11-11-2004